

## Appendix 1.

### The Distribution of OS with Crossovers

Let  $TTP \sim \text{Exp}(\lambda_1)$ ,  $X \sim \text{Exp}(\lambda_2)$ , where  $TTP$  is the time to progression and  $X$  is the survival time prior to progression. These distributions are the sub-distributions in a competing risks model. It follows that  $PFS = \min(TTP, X) \sim \text{Exp}(\lambda_1 + \lambda_2)$ . Furthermore, let  $OS' \sim \text{Exp}(\lambda_3)$  for patients who do not crossover and  $OS' \sim \text{Exp}(\lambda_4)$  for crossover patients, where  $OS'$  is post-progression survival time.

The probability density function for  $OS'$  is

$$f(OS') = p\lambda_4 \exp^{-\lambda_4 t} + (1 - p)\lambda_3 \exp^{-\lambda_3 t}$$

The probability distribution function for  $OS$  can be written as

$$\begin{aligned} F_{OS}(t) &= P(OS \leq t) \\ &= P(OS \leq t \mid OS = PFS)P(OS = PFS) + P(OS \leq t \mid OS > PFS)P(OS > PFS) \end{aligned}$$

where

$$\begin{aligned} P(OS \leq t \mid OS = PFS) &= P(PFS \leq t) = 1 - \exp^{-(\lambda_1 + \lambda_2)t} \\ P(OS = PFS) &= P(X \leq TTP) = \frac{\lambda_2}{\lambda_1 + \lambda_2} \\ P(OS > PFS) &= P(X > TTP) = \frac{\lambda_1}{\lambda_1 + \lambda_2} \\ P(OS \leq t \mid OS > PFS) &= P(PFS + OS' \leq t) = P(PFS \leq t - OS') \end{aligned}$$

The latter equation can be written as

$$\begin{aligned}
P(PFS \leq t - OS') &= \int_0^t \left[ \int_0^{t-v} (\lambda_1 + \lambda_2) \exp^{-(\lambda_1 + \lambda_2)u} du \right] [p\lambda_4 \exp^{-\lambda_4 v} + \lambda_3(1-p) \exp^{-\lambda_3 v}] dv \\
&= \int_0^t \left( 1 - \exp^{-(\lambda_1 + \lambda_2)(t-v)} \right) [p\lambda_4 \exp^{-\lambda_4 v} + \lambda_3(1-p) \exp^{-\lambda_3 v}] dv \\
&= \int_0^t \left[ p\lambda_4 \exp^{-\lambda_4 v} + \lambda_3(1-p) \exp^{-\lambda_3 v} - p\lambda_4 \exp^{-(\lambda_1 + \lambda_2)(t-v) - \lambda_4 v} \right. \\
&\quad \left. - \lambda_3(1-p) \exp^{-(\lambda_1 + \lambda_2)(t-v) - \lambda_3 v} \right] dv \\
&= p(1 - \exp^{-\lambda_4 t}) - (1-p)(\exp^{-\lambda_3 t} - 1) + \frac{p\lambda_4}{\lambda_1 + \lambda_2 - \lambda_4} \left( \exp^{-(\lambda_1 + \lambda_2)t} - \exp^{-\lambda_4 t} \right) \\
&\quad - \frac{(1-p)\lambda_3}{\lambda_1 + \lambda_2 - \lambda_3} \left( \exp^{-\lambda_3 t} - \exp^{-(\lambda_1 + \lambda_2)t} \right) \\
&= 1 - \frac{\lambda_1 + \lambda_2}{\lambda_1 + \lambda_2 - \lambda_4} p \exp^{-\lambda_4 t} - \frac{\lambda_1 + \lambda_2}{\lambda_1 + \lambda_2 - \lambda_3} (1-p) \exp^{-\lambda_3 t} \\
&\quad + \left( \frac{p\lambda_4}{\lambda_1 + \lambda_2 - \lambda_4} + \frac{(1-p)\lambda_3}{\lambda_1 + \lambda_2 - \lambda_3} \right) \exp^{-(\lambda_1 + \lambda_2)t}
\end{aligned}$$

Thus we can write

$$\begin{aligned}
F_{OS}(t) &= P(OS \leq t) \\
&= \left( 1 - \exp^{-(\lambda_1 + \lambda_2)t} \right) \frac{\lambda_2}{\lambda_1 + \lambda_2} \\
&\quad + \left[ 1 - \frac{\lambda_1 + \lambda_2}{\lambda_1 + \lambda_2 - \lambda_4} p \exp^{-\lambda_4 t} - \frac{\lambda_1 + \lambda_2}{\lambda_1 + \lambda_2 - \lambda_3} (1-p) \exp^{-\lambda_3 t} \right] \frac{\lambda_1}{\lambda_1 + \lambda_2} \\
&\quad + \left[ \left( \frac{p\lambda_4}{\lambda_1 + \lambda_2 - \lambda_4} + \frac{(1-p)\lambda_3}{\lambda_1 + \lambda_2 - \lambda_3} \right) \exp^{-(\lambda_1 + \lambda_2)t} \right] \frac{\lambda_1}{\lambda_1 + \lambda_2} \\
&= 1 - \frac{p\lambda_1 \exp^{-\lambda_4 t}}{\lambda_1 + \lambda_2 - \lambda_4} - \frac{(1-p)\lambda_1 \exp^{-\lambda_3 t}}{\lambda_1 + \lambda_2 - \lambda_3} \\
&\quad + \frac{\exp^{-(\lambda_1 + \lambda_2)t}}{\lambda_1 + \lambda_2} \left[ \frac{p\lambda_1 \lambda_4}{\lambda_1 + \lambda_2 - \lambda_4} + \frac{(1-p)\lambda_1 \lambda_3}{\lambda_1 + \lambda_2 - \lambda_3} - \lambda_2 \right]
\end{aligned}$$

The probability density function for  $OS$  is

$$f_{OS}(t) = \frac{p\lambda_1\lambda_4}{\lambda_1 + \lambda_2 - \lambda_4} \exp^{-\lambda_4 t} + \frac{(1-p)\lambda_1\lambda_3}{\lambda_1 + \lambda_2 - \lambda_3} \exp^{-\lambda_3 t} - \left[ \frac{p\lambda_1\lambda_4}{\lambda_1 + \lambda_2 - \lambda_4} + \frac{(1-p)\lambda_1\lambda_3}{\lambda_1 + \lambda_2 - \lambda_3} - \lambda_2 \right] \exp^{-(\lambda_1 + \lambda_2)t}$$

and the hazard function for  $OS$  is

$$\lambda_{OS}(t) = \frac{f_{OS}(t)}{S_{OS}(t)}$$

where  $S_{OS}(t) = 1 - F_{OS}(t)$ .

## Appendix 2

### Simulation Approach for Determining $D^*$

For each scenario and each iteration, 5,000 trials are simulated and a two-sample logrank test comparing the treatment arms with respect to OS is computed for each trial. The simulations were carried out using the R programming language (version 3.0.1). Since the required  $D^*$  in a non-proportional situation depends on the accrual time  $T$  and the accrual rate  $a$ , where  $N = aT$ , to determine  $D^*$  we start the iteration with an initial value  $D_0$  and an initial sample size  $N_0 = \gamma D_0$ , where  $\gamma \geq 1$  is pre-specified. Then we simulate the power  $1 - \beta_0$  for the  $(N_0, D_0)$  pair. For the next iteration,  $D_1 = \lceil D_0 + \epsilon_0 \rceil$ ,  $N_1 = \gamma D_1$ . Continue in this fashion, where for  $k \geq 0$  we define  $D_{k+1} = \lceil D_k + \epsilon_k \rceil$ ,  $N_{k+1} = \gamma D_{k+1}$ , with the starting value  $D_0$  and increments  $\epsilon_k$  defined below. The process stops as soon as  $|\beta - \beta_k| \leq \delta$ , a value that can be chosen by the user. We used  $\delta = 0.0025$  in this paper. That is, we stop at the first iteration for which the simulated power  $1 - \beta_k$  is within 0.0025 of the required power  $1 - \beta$ . The starting value  $D_0$  is:

$$D_0 = \left\lceil \frac{4(Z_{\alpha/2} + Z_\beta)^2}{(\ln \Delta)^2} \right\rceil$$

where  $\Delta$  is taken to be the ratio of the hypothesized median survival times in the two groups. This is the required number of events in the case of proportional hazards which of course does not hold here. The increments  $\epsilon_k$  are defined as:

$$\epsilon_k = \frac{4}{(\ln \Delta)^2} [(Z_{\alpha/2} + Z_\beta)^2 - (Z_{\alpha/2} + Z_{\beta_k})^2]$$

Note that  $\epsilon_k < 0$  if  $1 - \beta_k > 1 - \beta$  and  $\epsilon_k > 0$  if  $1 - \beta_k < 1 - \beta$ .

When the stopping rule is met, we perform an additional independent validation step using 10,000 simulations with the selected  $D_k$  and  $N_k$ . If the simulated power  $1 - \beta'_k$  in this additional step also satisfies  $|\beta - \beta'_k| \leq 0.0025$  then we stop and declare  $D^* = D_k$ . But if  $|\beta - \beta'_k| > 0.0025$  we continue the process

using  $\beta'_k$  in the next step. In practice we have found that the process stops after a very few iterations and the validation simulation almost always confirms the result without the need for additional iterations.

Example below illustrates the algorithm for determining  $D^*$  when  $p = 0.75$ . With the design parameters listed in Table 2 and  $\gamma = 1.1$ , the starting value  $D_0$  is

$$D_0 = \left\lceil \frac{4(Z_{\alpha/2} + Z_{\beta})^2}{(\ln \Delta)^2} \right\rceil = 818$$

and the initial sample size  $N_0 = 1.1D_0 = 900$ . The simulated power  $1 - \beta_0$  for the the  $N_0$  and  $D_0$  pair is 0.8526. Since  $|\beta - \beta_0| > 0.0025$ , we continue the process and the increments  $\epsilon_0$  is

$$\epsilon_0 = \frac{4}{(\ln \Delta)^2} [(Z_{\alpha/2} + Z_{\beta})^2 - (Z_{\alpha/2} + Z_{\beta_0})^2] = 113.692$$

For the next iteration, we have  $D_1 = \lceil D_0 + \epsilon_0 \rceil = 932$ ,  $N_1 = \gamma D_1 = 1025$ . The results of subsequent iterations are given in Table 5.

Table 5: Results for Determining  $D^*$  When  $p = 0.75$

$k$	$N_k$	$D_k$	$1 - \beta_k$	$ \beta - \beta_k $	$\epsilon_k$
0	900	818	0.8526	0.0474	113.692
1	1025	932	0.8948	0.0052	14.6007
2	1042	947	0.8980	0.0020	5.6952

To validate the selected  $D_2 = 947$  and  $N_2 = 1042$ , an independent step using 10,000 simulations is performed. The simulated power  $1 - \beta'_2 = 0.8984$  satisfies the  $|\beta - \beta'_2| \leq 0.0025$  requirement, which validates the selected  $D_2 = 947$ . We stop and declare that  $D^* = D_2 = 947$  for  $p = 0.75$ . The R code for determining  $D^*$  is also provided as supplemental material.